Novel peptide replicators from dynamic combinatorial libraries

Altay, Yigit

DOI:
10.33612/diss.90041906

Chapter 1

Introduction

"*Sane sicut lux seipsam et tenebras manifestat, sic veritas norma sui et falsi est*" [1]

Spinoza

Abstract

How we define life is fundamental to the process of understanding it. Since ancient times, it has been a puzzling subject to philosophers, writers and scientists. In the last 150 years, the number of available tools in the hands of scientists to probe this question increased almost exponentially. Now we are on the edge of understanding the mechanisms of life and possibly creation of an alternative lifelike form by our hands.

This chapter introduces systems chemistry and dynamic combinatorial chemistry as useful tools to shed light on the complex molecular systems that exhibit emergent properties that are analogous to those found in nature.

1.1 What is life?

What is life? This grammatically very simple question is one of the hardest questions to yield in an operational and complete answer. Historically, there are many attempts made to define life[2–7] and to distinguish between what is living and what is non-living. A sharp division exists in most cases and we call complex crystals as non-living and rather more complex animals as living. However, in the last century, science provided evidence that there is a continuum between the two. A recent definition of life is given by NASA as "a self-sustaining chemical system capable of Darwinian evolution". This broad definition is far from being complete but has the advantage of including a wide range of potential life forms.

Even though there might not be a complete definition, life as we know has certain characteristics that may help us identify it. First, it needs to have a body to contain the self and distinguish it from the environment. Formation of the body can be considered as compartmentalization. In this confined environment, an organized chemical system is contained. Second, there must be a process (or a series of processes that we can call metabolism) that converts raw materials into energy and components to sustain itself. Through simple or more complex metabolic reactions, the body can provide energy and products that can be processed. As a result, it can start to interact with its environment. And last, there must be a transfer (inheritance) of the basic data that defines that particular body over generations through reproduction. The transfer of data does not necessarily need to be perfect but might contain imperfections. As a result of these imperfections, the body can undergo Darwinian evolution.

These characteristics allow us to study and understand the mechanisms of life and possibly create life *de novo* from purely synthetic parts. In contrast to the biological perspective, where a top-down approach is applied, chemists prefer to use bottom-up approaches to deal with the same problem. The main advantage of the bottom-up approach is that it does not limit the tools to the ones that are provided by the nature. Although the main source of inspiration is still nature itself, synthetic chemistry offers a wide range of possibilities to use in the search for the origin of life and in creating life *de novo*.

1.2 Systems Chemistry

Biology studies natural systems in their entire complexity whereas traditional chemistry deals with pure compounds. It is clear that traditional chemistry cannot probe the complexity of nature by such focus on isolated molecules and thus bridging the gap between chemistry and biology requires a new approach. Advancements started

in the late 1960s with the rise of supramolecular chemistry[8] which studies assemblies formed through the non-covalent interactions between molecules. Afterwards, systems chemistry stepped in by studying complex molecular networks that interact with each other and giving rise to emergent properties.[9–16] These properties are resulting from the collective behavior in the system and they cannot be ascribed to any of the individual components. In 2007, a new European Union research network (COST CM0703) defined the systems chemistry as:[17]

> " ... the joint effort of prebiotic and supramolecular chemistry assisted by computer science from theoretical chemistry, biology, and complex systems research to tackle dynamic supersystem integration including at least one autocatalytic subsystem. It is the bottom-up pendant of systems biology towards synthetic biology. The origin of life is seen as a major stimulus to organize research but the field is open for chemistries of limited prebiotic plausibility. Subsystems may be classified as genetic, metabolic, or compartment-building. Pairwise integration into higher organized supersystems is expected to yield the knowledge enabling later the triple integration into minimal chemical cells. The integration approach will necessarily link to the question of asymmetric autocatalysis and chiral symmetry breaking, while the key challenge is to find the roots of Darwinian evolvability in chemical systems."

Initial research area of systems chemistry centered around protocell research[18,19] and started to attract more attention after the inspiring work of Szostak *et al.*[20] on how to synthesize life. Hereon, the systems chemistry approach started to be used in the search of new biochemistries. This includes a wide range of research areas from the design of nucleic acid alternatives[21–23] to enzyme-free molecular networks.[24]

1.3 Dynamic Combinatorial Chemistry

Dynamic combinatorial chemistry (DCC)[15,25–31] is an important tool in system chemistry as it allows the facile construction of complex molecular networks in which emergent properties can be investigated through the response of the library to the external stimuli. Dynamic combinatorial libraries (DCLs)[32] are composed of building blocks that are functionalized with groups that allow reversible bond formation. The reversibility is the main source of the dynamic nature of these systems and it provides the foundation of the network of interconverting compounds. Under thermodynamic control, the library composition is dictated by the relative free energies of each library member and the self-assembled structures. However, any external stimulus that can alter the relative stability of a library member can effect the library distribution.

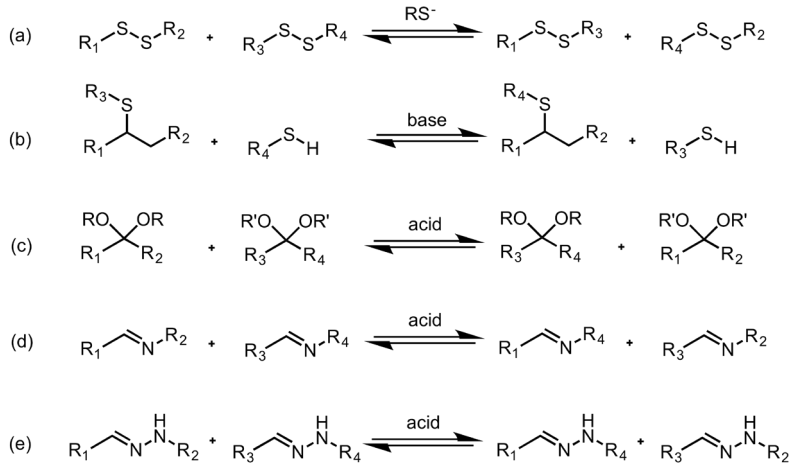

Figure 1.1: Common reversible reactions used in dynamic combinatorial libraries: a) disulfide exchange, b) Michael / retro-Michael reactions, c) acetal exchange, d) imine exchange and e) hydrazone exchange.

Among the various exchange reactions (see Figure 1.1 for the most common ones), disulfide exchange is one of the most commonly used reactions in DCC as it provides good control over the exchange process. The disulfide exchange reaction runs under relatively mild conditions. In slightly basic conditions (pH around 8) thiol-containing building blocks form disulfides upon oxidation by the oxygen present in air (Figure 1.2a). In the presence of thiolate anions, the exchange reaction takes place (Figure 1.2b). After all the thiols are oxidized or when they are protonated, exchange is no longer possible. Reduction of the system using reducing agents such as dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP) converts disulfides into thiols and the system becomes dynamic again. The disulfide exchange reaction is also highly tolerant to different functional groups.[33]

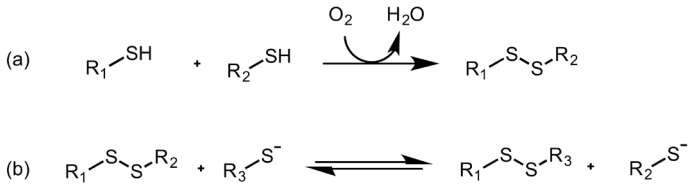

Figure 1.2: a) Disulfide bond formation upon oxidation and b) mechanism of disulfide exchange in the presence of thiolate ion.

1.4 Template Induced Self-Assembly

In 1924, Russian chemist Alexander Oparin proposed that life on earth emerged spontaneously from inanimate matter via gradual increase in complexity.[34,35] Supramolecular chemistry, followed by the rise of systems chemistry, brought new perspectives on interaction space in chemistry that allows exploration of self-assembling chemical systems. Self-assembly and self-organization are the key concepts. In some cases they are erroneously used as synonyms, but they actually represent two sides of the same coin.[36] Although both of these processes are resulting from non-covalent interactions (such as H-bonding, dipole and van der Waals interactions etc.), they have a different thermodynamic basis. Self-assembly relies on spontaneous processes that are leading to equilibrium. In contrast, self-organisation generally implies dynamic non-equilibrium processes that result in ordered structures at nano- or micro-scale.

In dynamic combinatorial libraries, the product distribution of a thermodynamically controlled system can be altered via templates. Template molecule can either be a building block of the library or can be produced by the library itself (internal templating) or it can be a molecule that interacts with the components of the library and facilitate bond formations (external templating). In either case, binding to the template will result in amplification of the bound library member. This phenomenon is exploited mainly in the area of host-guest chemistry and in drug discovery as it allows to design and develop molecular recognition systems.

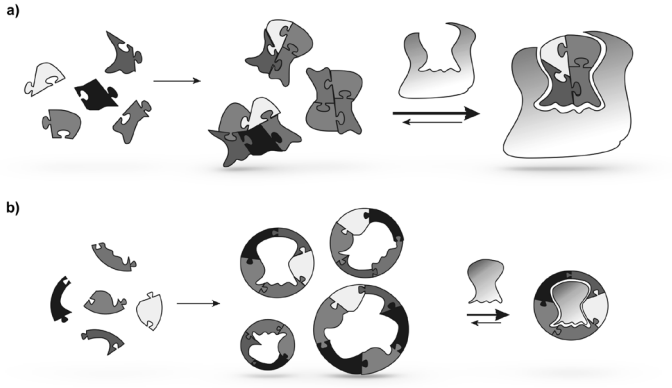

Figure 1.3: Selection and amplification of a library member by external templating via a) a host molecule and b) a guest molecule. Figure reproduced from the reference.[37]

Metal coordination cages that are developed by using such host-guest interactions are very popular, among other supramolecular architectures. In this area, the

Nitschke group shows a state of the art advancement of such systems. In a recent work,[38] Black *et al.* developed a dynamic combinatorial library of Zn_4L_6 cages made from naphthalene diimide (NDI, **N**) and/or Zn-porphyrin moieties (**P**) (Figure 1.4) where they showed an example of the modulation of guest inclusion within a complex chemical network. Similar lengths of the cage edges allow formation of mixed species whereas the different shapes of **N** and **P** moieties results in specific host-guest interactions. Consequently, the DCL composed of 7 different cages responds differently upon addition of different templates. Addition of C_{70} molecule to the DCL shifts the equilibrium towards to C_{70} encapsulated $\mathbf{P}_6$ cages resulting in a self-sorting behavior towards a homoleptic cage. On the other hand, addition of bis-1,5-(dinaphtho)-38-crown-10 (**C**) results in selective amplification of heteroleptic cages via catenation of electron rich **C** and electron poor **N** moieties. Due to steric reasons more than 2 **C** cannot catenate on a cage. This work is a nice demonstration of external templation where the templation can induce two different behaviors in the same library.

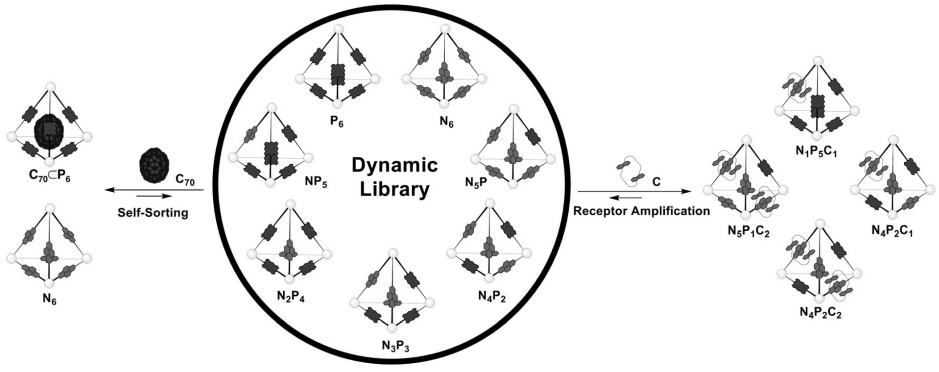

Figure 1.4: Figure adapted from the reference.[38]

For internal templation, the Stoddart group reports an excellent example of template directed synthesis of oligo-rotaxanes that is making use of the DCC approach.[39] Figure 1.5 shows the building blocks **1a-f**·$n$$PF_6$ containing dibenzylammonium centres functionalized with two dimethoxybenzyl groups at each end, 2,6-pyridinedicarboxaldehyde (**2a**) and tetraethyleneglycolbis(2-aminophenyl)ether (**3**). Mixing all three building blocks in equimolar concentrations results in quantitative formation of [3]-, [4]-, [5]-, and [7]rotaxanes (**4a-4d**·$n$$PF_6$) in a few minutes. The ring formation reaction between **2** and **3** is templated by **1a-1d**·$n$$PF_6$. Non-covalent interactions, particularly π-π stacking between the thread and [24]crown-8 diimine-containing rings and H-bonding between [N^+-H$\cdots$X] and [N^+-C-H$\cdots$X] (where X=O or N), are driving forces for this reaction to occur in the DCC. For the formation of

higher order oligorotaxanes, [11]- and [15]-rotaxanes, 4-octyloxy-functionalized 2,6-pyridinedicarboxaldehyde (**2b**) was used to increase the solubility of the corresponding poorly soluble threads. These kinetically stable rotaxanes can be fixed by reduction of the dynamic imine bonds.

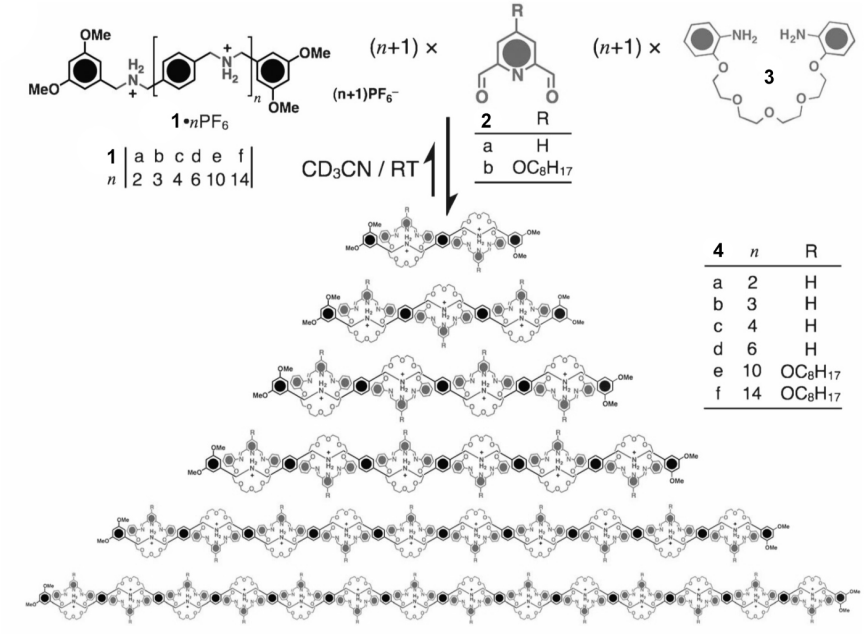

Figure 1.5: Template-directed synthesis of the dynamic oligorotaxanes **4a-4f**·nPF$_6$ from simple building blocks **1a-f**·nPF$_6$, **2a-b** and **3**. Figure adapted from the reference.[40]

1.5 Self-Replication

Self-replication is an indispensable element of life since it is the main source of propagation of life itself. Replication is often an imperfect process in most organisms that allows natural selection[41] so it becomes an intrinsic requirement for Darwinian evolution where organisms form closely-related species.[42]

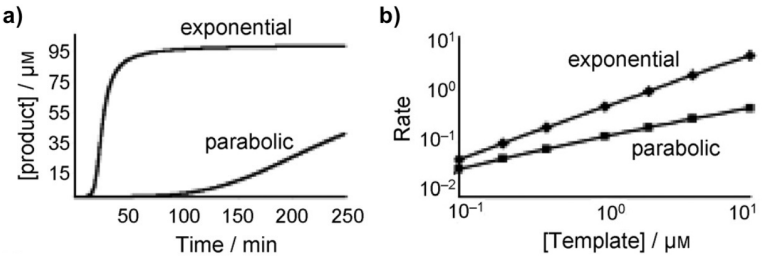

Figure 1.6: Kinetic growth curve of an autocatalytic reaction. a) Concentration vs. time and b) rate vs. initial concentration of product graph for exponential and parabolic growth. Figure adapted from reference.[24]

Self-replication is a subset of autocatalytic reactions in which the product acts as a catalyst for its own formation and in which (sequence) information is inherited. An autocatalytic reaction has a characteristic growth curve where the rate of the reaction is correlated to the initial concentration of the product (Figure 1.6). If the autocatalysis is not efficient, the growth curve shows parabolic behavior rather than exponential. In a self-replicating system in which the initial concentration of the product is zero, the growth curve becomes sigmoidal. Figure 1.7 shows such sigmoidal growth curve featuring three distinct zones. The lag phase ends when the autocatalytic cycle starts to produce enough products to catalyze its own formation. As the cycle produces more product, rate of reaction increases until the starting reagents start to run out. In the last phase, diminishing starting materials become the limiting factor and the rate of the reaction decreases and the growth curve reaches a plateau.

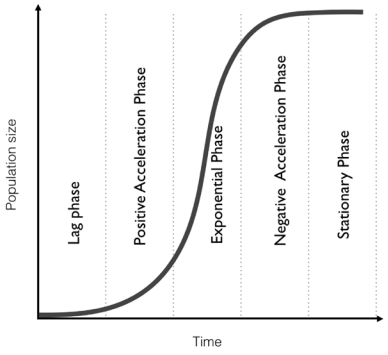

Figure 1.7: A representative plot of a sigmoidal growth curve featuring lag, exponential and stationary phases.

The minimal self-replication cycle[43–45] in Figure 1.8 shows components **A** and **B** with complementary functional groups and reactive ends. **A** and **B** react through an uncatalysed bimolecular reaction to form template molecule **T**. Similarly, **T** can be formed via an inactive form of **T** through a reaction of **A** and **B** after formation of a duplex [**A**·**B**]. In the autocatalytic cycle, **A** and **B** simultaneously bind to the template molecule **T** and form a ternary complex [**A**·**B**·**T**]. The relative reaction rate of **A** and **B** in the ternary complex increases due to programmed orientation resulting from the specific interactions between the template molecule and its precursors and the higher local concentration of the building blocks. Consequently, **A** and **B** react with each other to form a copy of **T**. In the last step of the autocatalytic cycle, duplex [**T**·**T**] dissociates to form two identical template molecules. At the end of n number of autocatalytic cycles, 2^n number of templates will be produced. However, the disassociation is often the rate determining step of an autocatalytic cycle as a result of the relatively strong interaction between two template molecules in the duplex. It should be noted that an efficient self-replicating system should not allow side reactions or at least, side reactions should be minimized.

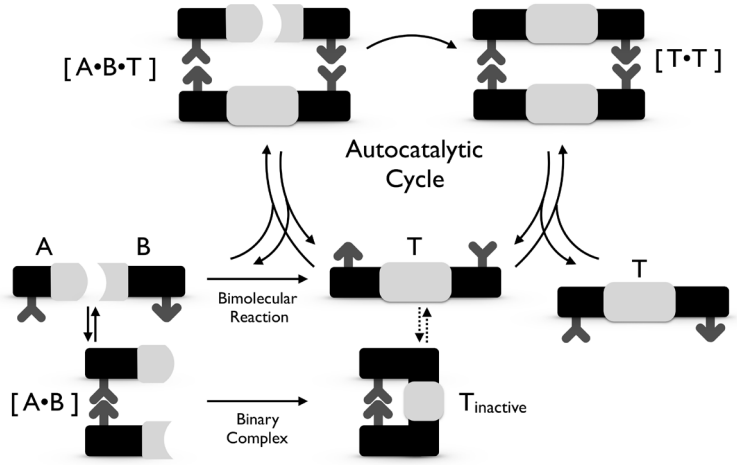

Figure 1.8: Minimal self-replicating system.

The design of self-replicating systems was inspired by nature. Hence, the first examples of the minimal self-replicating systems featured natural building blocks. In 1986, von Kiedrowski[46] reported the first example of minimal self-replicating system based on a hexadeoxynucleotide with a palindromic base sequence (Figure 1.9). His

system uses a strong base pairing interaction of cytosine-guanine between the building blocks of **A** with a sequence 5′-CCG-3′ and **B** with a sequence of 5′-CGG-3′. The 5′ terminus of building block **A** is protected by a methyl ether, while 3′ terminus contains a phosphate group. The building block **B**, on the other hand, is protected at the 3′ terminus by *o*-chlorophenyl to prevent ligation at this site whilst the 5′ terminus has a free hydroxy group that can react with a phosphate. In the ternary complex, the phosphate group at the 3′ terminus of **A** comes in close proximity to the hydroxyl group at the 5′ terminus of **B**. Thus, the ligation reaction is catalyzed by the template after the phosphate group is activated by 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide (CDI). Due to the palindromic base sequence of the template molecule, the reaction between **A** and **B** yields the exact same sequence 5′-CCGCGG-3′.

Unfortunately this system did not show a sigmoidal growth curve due to several competing reactions. However, von Kiedrowski proved the autocatalytic behavior of the system by showing that the initial rate of the reaction increased upon addition of small amount of template molecule. The system suffers from product inhibition possibly due to the hampered disassociation of the duplex [**T**·**T**].

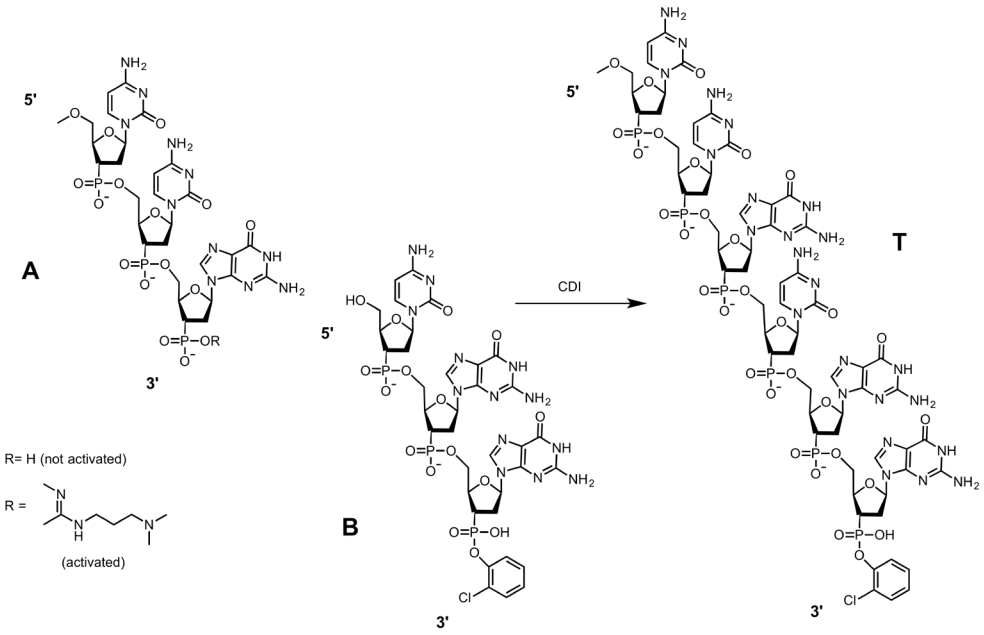

Figure 1.9: Self-replicating hexadeoxynucleotide reported by von Kiedrowski.

Following the pioneering work of von Kiedrowski, other nature-inspired non-

enzymatic systems have been developed with organic molecules, peptides, fatty acids and nucleic acids.[47–52] In 2012, the Ashkenasy group reported a self-replicating system based on a self-replicating peptide with its electrophilic and nucleophilic precursors.[53] Peptide (**5** in Figure 1.10) is made of a chain of alternating hydrophilic Glu and hydrophobic Phe residues on the sides. There is also a ligation site in the middle made from Cys-Ala and the Pro residue at the C-terminus to induce one-dimensional aggregation. As the system grows, temporary binding of precursors on the edges of the one dimensional structures made from peptide **5** facilitates the reaction between **E** and **N**. Thus, the structure can continue to grow by catalyzing the formation of its building blocks and then binding them on the edges. It is shown that monomeric peptide **5** cannot bind to precursors so that β-sheet nanostructures of **5** are crucial for the self-replication. β-sheet platelets forms helical β-sheet fibrils that are stable in solution only for a few hours because they assemble further into inactive hollow nanotubes. Overall, system catalyzes its own formation only through the transient nanostructures and after equilibration into nanotubular structures, the system becomes inactive.

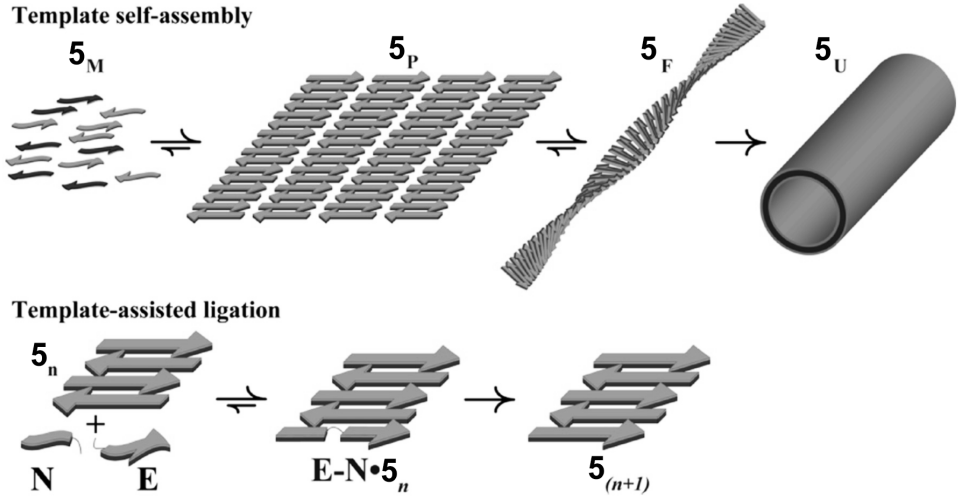

Figure 1.10: Schematic representation of simultaneous processes (template self-assembly and template assisted ligation) that, together, give rise to self-replication in Ashkenasy's system. Figure adapted from reference.[53]

The laboratory of Douglas Philp exploited and excelled in the area of minimal self-replicating systems using Diels-Alder[54–59] and 1,3-dipolar cycloaddition[60,61] reactions in the ligation step. In a recent study,[62] Robertson *et al.* demonstrate im-

plementation of reciprocal replication made from a complementary pair of templates using 1,3-dipolar cycloaddition and imine condensation reactions. Orthogonality between these reactions is the key element for this system to be successful and it allows the formation of the two cross-catalytic cycles, one containing **A**-**B** and its reciprocal partner template $\mathbf{T}^{CD}$ and the other from **C**-**D** and its reciprocal partner template $\mathbf{T}^{AB}$. When all four building blocks that form the two templates, $\mathbf{T}^{AB}$ and $\mathbf{T}^{CD}$, are mixed, an efficient self-replicating system can be formed and sustained. In the presence of reciprocal partners, the rate of formation of each template is significantly accelerated with respect to the rate of formation in isolated systems. Additionally, it is shown that the system can be directed by introduction of one of the two templates in the presence of all four building blocks. In such a case, only the production of the reciprocal partner and the template added enhances. This system opens up new possibilities in the design of responsive dynamic self-replicating networks that can be controlled by the input of templates in addition to environmental factors.

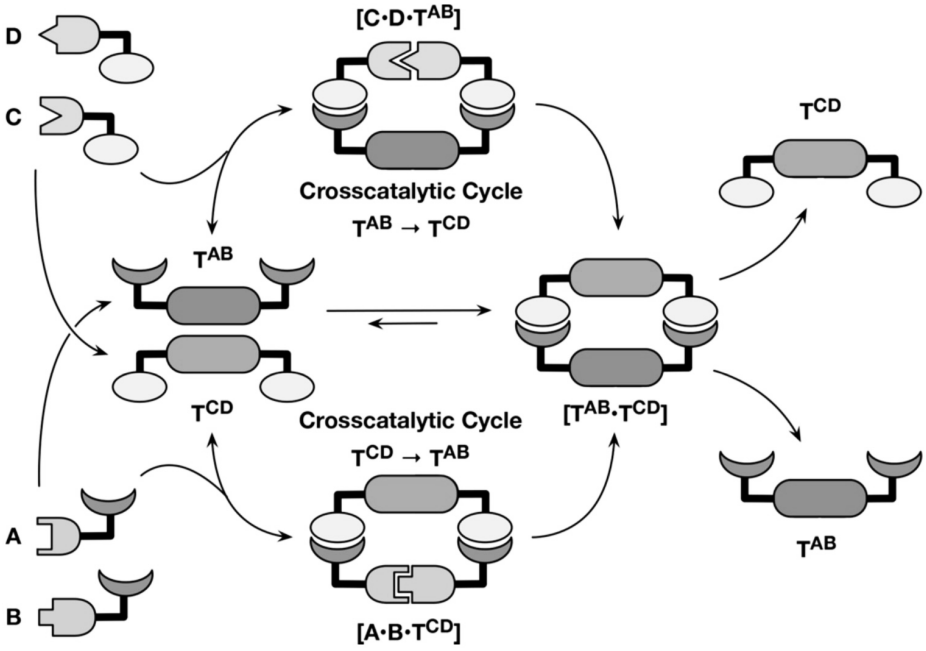

Figure 1.11: Schematic representation of a reciprocal self-replication cycle in a four building block system developed in the Philp laboratory. Figure adapted from reference.[62]

In mid-1990s, researchers started to incorporate simple self-replicating systems

into more complex networks where different (auto)catalytic cycles operates simultaneously. In 1998, Severin *et al.* reported a dynamic network that can correct errors (Figure 1.12) based on a replicating peptide system that was previously published.[63,64] Native electrophilic (**E**) and nucleophilic (**N**) fragments and their mutants (single alanine on positions 9 and 26) can form native peptide **T**, single mutants $\mathbf{T}_{9A}$ and $\mathbf{T}_{26A}$, and a double mutant $\mathbf{T}_{9A}/\mathbf{T}_{26A}$ via chemical ligation reactions. In the system at neutral pH, production of native peptide **T** is amplified as it catalyzes only its own formation. On the other hand, single mutants $\mathbf{T}_{9A}$ and $\mathbf{T}_{26A}$ show cross-catalytic activity in favor of formation of error-free peptide **T**. The double mutant $\mathbf{T}_{9A}/\mathbf{T}_{26A}$ is found to be catalytically inactive. The system bears two cross-catalytic cycles, **2** and **3** where the mutant peptides catalyze the formation of error-free peptide **T** operating simultaneously with the autocatalytic cycle **1** where peptide **T** selfishly catalyzes its own formation. In overall, the system can handle sequence specific self-replication and error-correction simultaneously.

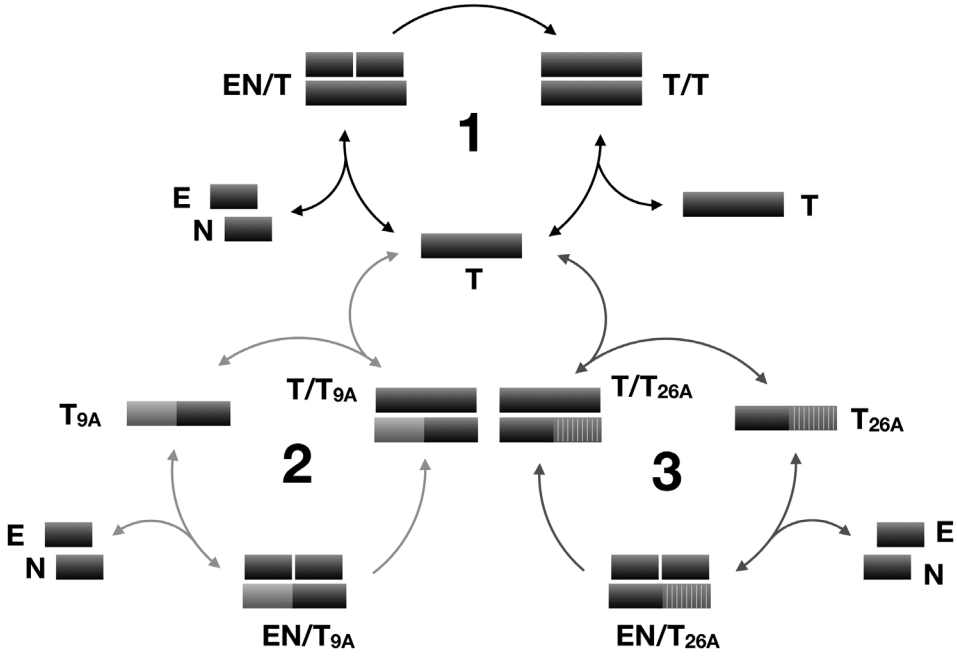

Figure 1.12: Error correcting self-replication system developed by Severin *et al.*

In 2010, a novel self-replicating peptide system[65] was discovered by our group and since then extensive studies on the system reveal new possibilities and pathways in understanding the nature of self-replication and the creation of life *de novo*. Figure

1.13 shows the schematic representation of our system that is composed of pentapeptide building blocks functionalized with an aromatic dithiol. The alternating hydrophobic and hydrophilic residues on the backbone facilitate interaction between the peptides and consequently drives β-sheet formation.

A peptide building block that does not have the aromatic dithiol functionality results in formation of a product mixture that is controlled by thermodynamics. Building blocks with thiols react at slightly basic conditions, around pH 8.0, to form intermolecular disulfide bonds which consequently form macrocycles that can exchange building blocks as long as there is a stock of building blocks in the solution. The size of the macrocycle is influenced by the strength of the interactions between the peptide building blocks and thus by the sequence of the peptide. This phenomenon was discovered previously[66] and is further elaborated in this thesis. For a detailed description of our system, see Chapter 2.

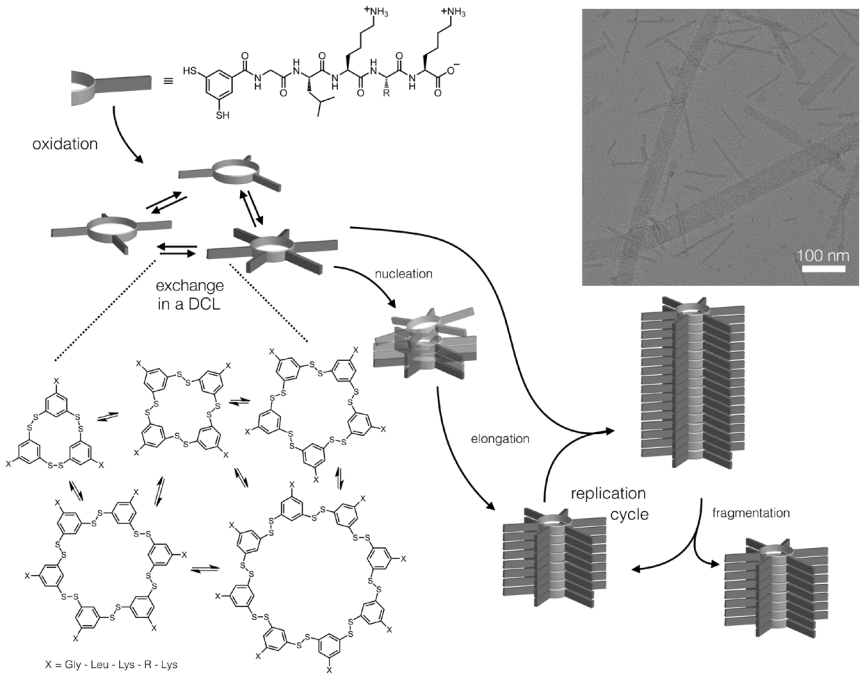

Figure 1.13: Schematic representation of self-replicating system developed in the group of Sijbren Otto.

1.6 Out of Equilibrium Systems and Quasi-Speciation

Supramolecular chemistry mostly focuses on systems at equilibrium, mainly because most systems are under thermodynamic control. Recently, kinetically controlled self-assembly processes, where the outcome is dictated by the pathway of the assembly, instead of the Gibbs free energy of the final state, are attracting attention. Such processes result in structurally and functionally more diversified assembly systems. Life itself is one of the best examples illustrating the richness of systems under kinetic control. Fundamental processes of Darwinian evolution, namely selection and adaptation, are driven by the continuous formation and destruction of the (bio)molecular assemblies. Stability of these assemblies heavily depends on the adaptation to the environments that are out of equilibrium rather than to thermodynamically controlled processes.[42,67,68] So, exploration in this direction would open up new perspectives in understanding the basics of life and possibly creating life *de novo*.

Some out-of-equilibrium systems require a continuous supply of energy to be sustained; i.e. they will persist only as long as energy is provided. As the energy supply is continued, transformations that lead to often unpredictable emergent functions can appear in the system. As soon as the energy supply stops, the system collapses and ends up in the thermodynamically most stable state (alternatively it can end up in a kinetic trap if it exists on the pathway to equilibrium). Experimentally, it is not easy to probe an out-of-equilibrium system and commonly systems are studied as they are approaching equilibrium state. The Belousov-Zhabotinsky (BZ) reaction[69,70] is a classical example of such non-equilibrium system showing oscillations in the concentrations of some of the species in the reaction network over time. In a closed system also the BZ reaction would eventually fall into an equilibrium state (or a kinetic trap) and the emergent properties (oscillations) that arose initially do not persist. In order to sustain these emergent properties, the system should be kept in a steady state where a continuous inflow is synchronized with an outflow.

Recently, synthetic systems showing out-of-equilibrium self-assembly are attracting increasing interest. An example of non-equilibrium biocatalytic gelation of peptide nanofibers is shown in Figure 1.14. Self-assembly of peptide nanostructures that are resulting from competitive assembly and disassembly reactions display dynamic instability in which the assembly is favored away from equilibrium.[71] The forward reaction is a kinetically controlled trans-acylation catalyzed by α-chymotrypsin. The reverse reaction is a thermodynamically controlled disassembly reaction based on amide hydrolysis. The system uses the chemical energy stored in the methyl ester precursor for gelation and thus creates a temporary high concentration. Hydrolysis results in the formation of nanofibers that have dynamic instability. Importantly, the system can be refueled up to three times. After that point *in situ* reactivation

is not possible due to accumulation of Nap-Y-OH and refueling attempts by methyl iodide failed as the reaction is not selective. Overall, this work represents a mimicry of natural dynamic systems based on much simpler building blocks. An analogous system has also been reported by the van Esch group.[72]

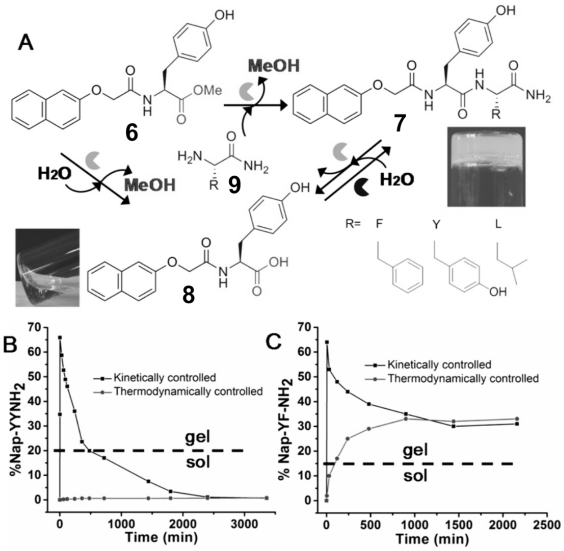

Figure 1.14: Non-equilibrium biocatalytic self-assembly system by Ulijn's group. Figure adapted from reference.[71]

A second example deals with far-from-equilibrium self-assembly of nanoparticles that represents rich complex behavior. In recent work of Ilday *et al.*,[73] tens to thousands of nanoparticles exhibit complex behavior that includes autocatalysis, self-replication, self-regulation, competition, adaptation and self-healing. The system is triggered by the ultrafast laser pulses that create spatiotemporal temperature gradients which drives aggregation induced by Marangoni-type microfluidic flow, while strong Brownian motion counteracts the aggregation. As illustrated in Figure 1.15a, the system does not include functionalized nanoparticles or any commonly used interaction tools such as chemical or magnetic interaction, tweezing or optical trapping. Instead, it relies on (1) nonlinearity to create multiple steady states that each represents a different pattern and their bifurcations, (2) exponential growth of perturbations that is resulting from positive and negative feedback loops, (3) spontaneous transitions between bifurcations via fluctuations and (4) spatiotemporal temperature gradients to operate the system far from equilibrium and create coexistence and dynamic growth/destruction within regions. The agreement between a toy model

(Figure 1.15f) and experimental data supports the autocatalytic behavior of the system. When the laser was turned off, as the particles are in thermal equilibrium, normal fluctuations resulting from Brownian motion are observed. After the laser was turned on, nanoparticles are accelerated and dragged by the Marangoni-type microfluidic flow.

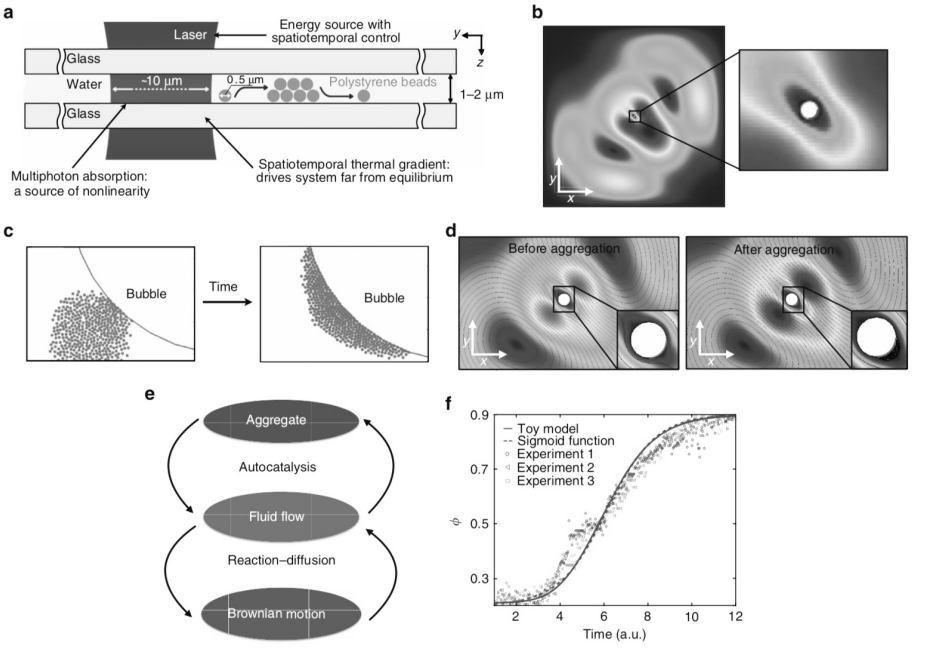

Figure 1.15: Assembling nanoparticles are showing rich complex behavior far from equilibrium. a) Illustration of the experimental setup: polystyrene beads sandwiched between two glass plates with a ultrafast laser. b) Velocity field simulation of Marangoni-type microfluidic flow, c) Numerical simulation of the Brownian nanoparticles, d) Velocity field simulations of the flow before and after an aggregate forms, e) Schematic description of the nonlinear feedback mechanisms, f) Plot comparing toy model and three measurements showing a sigmoidal behavior. Figure adapted from reference.[73]

In thermodynamic processes, rare events might not play a significant role but those rare events can become crucial if they are allowed to amplify in an autocatalytic process. As we mentioned earlier, out-of-equilibrium conditions also increase the probability of creation of such rare events. Thus autocatalytic systems operating under out-of-equilibrium conditions have significant importance in molecular quasi-speciation. In biology, quasi-species are defined as a well-defined set of genotypes in a

population that is generated by mutation-selection process.[74,75] Virus populations in general exhibit extensive mutant distributions and so, recently, they have been called quasi-species to express this strong genetic heterogeneity.[75] At the molecular level, quasi-speciation refers to a distribution of molecules that are undergoing adaptations and (dynamic) transformations (analogous to mutations), where the resulting set of molecular distribution is capable of sustaining itself. A peptide-based self-replicator system developed by Sadownik *et al.*[76] represents a promising starting point for generating a quasi-speciation system where a set of replicator families emerges from a mixed building block system. In Chapter 5, a detailed explanation of the system is provided, together with further steps necessary to achieve quasi-speciation in this system.

1.7 Conclusion and Contents of this Thesis

Systems chemistry, and dynamic combinatorial chemistry in particular, are promising approaches to study complex molecular systems to answer fundamental questions at the interface between biology and chemistry. This thesis focuses on the exploration of novel self-replicating peptides from dynamic combinatorial libraries using reversible disulfide chemistry. In addition to structural exploration, interactions of different replicators are explored to access new pathways that lead to the emergence of such self-replicating systems.

In **Chapter 2**, we conducted a survey of self-replicating peptides emerging from dynamic combinatorial libraries. Novel peptide sequences are designed and synthesized to analyze the effect of minor structural variations on the emergence of self-replicators. C-termini modifications and salt effects were also investigated to understand the mechanism and the requirements of such systems. Results revealed that the hydrophobicity of the peptide is dominant up to certain level and beyond that, several other factors become more important. We also showed that the rate of formation and morphology of the self-assembled fibers are effected by the ionic strength of the buffers used.

We investigated the cross-catalytic relation between the threonine-containing peptide building blocks and readily formed replicator seeds made from different building blocks in **Chapter 3**. Macrocycles formed from the threonine peptide do not show any self-replication behavior by themselves; only non-assembling trimers and tetramers are formed. We showed the emergence of a new hexamer replicator made purely from threonine peptides upon seeding by a different replicator that has a specific ring size and has a specific side chain interaction. We also found that the history of the library affects the emergence of the replicator. By these results, we showed for the first time how a novel replicator can emerge with the assistance of another

replicator.

Encouraged by these results, we further explored how the presence of different replicators can direct the emergence of new replicators in **Chapter 4**. We selected a tyrosine-containing peptide building block for this study as the preliminary results, which are described in Chapter 1, showed that the energy profile of the system that is made from the tyrosine-peptide would allow us to control the outcome. Upon seeding by hexamer or octamer replicators, libraries produced replicators where the size of the emerging replicator is dictated by the size of the seed. For the octamer replicator, we also showed that not only the ring size but also the fiber morphology can be inherited.

In **Chapter 5**, we introduced a second building block to the system. The resulting family of macrocycles that are made from mixed building blocks showed a Gaussian distribution and is a good candidate to study quasi-speciation in completely synthetic systems. First, we ran several optimizations on the buffer preparation and on the experimental setup as well as on the infusion and withdrawal flow rate. Sustaining a replicator distribution over a few generations, where both formation and destruction mechanisms are enabled, was achieved using a mixed system made from phenylalanine- and serine-based peptide building blocks.

In **Chapter 6**, we introduced a new way of analyzing replicator composition in dynamic combinatorial libraries using a combinatorial fluorescent molecular sensor. As we embarked on new directions in our research, we realized that the statistical analysis of DCLs is becoming more and more important. For such analyses, the time required for chromatography becomes a major bottleneck. We showed that we can correlate the compositions of libraries containing self-assembled replicators as well as non-assembled macrocycles to optical fingerprints. We proved that we can distinguish not only different sized replicators but also the same sized replicators with minor structural variations. Lastly, we showed that we can track the emergence of a hexamer replicator on a model system. These results will allow researchers to study large numbers of sample sets in a short time with a smaller sampling volume and concentration.

Finally, **Chapter 7** gives an overview of this thesis and places the results and their implications in a broader context.

1.8 Acknowledgements

I. Maric is gratefully acknowledged for proof-reading this chapter and for giving precious feedback.

1.9 References

[1] Translation: "Truly, as light manifests itself and darkness, thus truth is the standard of itself and of error", *Ethics.* Pt II, Prop. 43.

[2] Allan, D. J. *The Philosophy of the Aristotle;* Oxford University Press: London, 1970.

[3] Farley, J. *The spontaneous generation controversy;* John Hopkins University Press: Baltimore, 1977.

[4] Mayr, E. *The growth of biological thought;* Harvard University Press: Cambridge, 1982.

[5] Mayr, E. *Toward a new philosophy of biology: Observations of an evolutionist;* The Belknap Press of Harvard University Press: Cambridge, 1988.

[6] Küppers, B. O. *Information and the origin of life;* MIT Press: Cambridge, 1990.

[7] Magner, L. N. *A history of the life sciences;* Marcel Dekker, Inc.: New York, 2nd edition ed.; 1994.

[8] Pedersen, C. J. *J. Am. Chem. Soc.* **1967,** *89,* 2495-2496.

[9] Gibb, B. C. *Beilstein J. Org. Chem.* **2016,** *12,* 684-701.

[10] Mattia, E.; Otto, S. *Nat. Nanotechnol.* **2015,** *10,* 111–119.

[11] Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. *Chem. Rev.* **2014,** *114,* 285-366.

[12] Stoddart, J. F. *Angew. Chem. Int. Ed.* **2012,** *51,* 12902-12903.

[13] Nitschke, J. R. *Nature* **2009,** *462,* 736-738.

[14] Peyralans, J. J. P.; Otto, S. *Curr. Opin. in Chem. Biol.* **2009,** *13,* 705 - 713.

[15] Ludlow, R. F.; Otto, S. *Chem. Soc. Rev.* **2008,** *37,* 101-108.

[16] Whitesides, G.; Ismagilov, R. *Science* **1999,** *284,* 89-92.

[17] Kiedrowski, G.; Otto, S.; Herdewijn, P. *J. Syst. Chem.* **2010,** *1,* 1.

[18] Luisi, P. L. *The Emergence of Life: From Chemical Origins to Synthetic Biology;* Cambridge University Press: Cambridge, 2006.

[19] Rasmussen, S.; Bedau, M.; Chen, L.; Deamer, D.; Krakauer, D. C.; Packard, N.; Stadler, P. F. *Protocells: Bridging Nonliving and Living Matter;* MIT Press: Cambridge, 2008.

[20] Szostak, J. W.; Bartel, D. P.; Luisi, P. L. *Nature* **2001,** *409,* 387.

[21] Herdewijn, P.; Marlière, P. *Chem. Biodivers.* **2009,** *6,* 791-808.

[22] Böhler, C.; Nielsen, P.; Orgel, L. *Nature* **1995,** *367,* 578-581.

[23] Nielsen, P. E. Peptide Nucleic Acids and the origin of life. In *Origin of Life: Chemical approach*; Herdewijn, P.; Kisakurek, V., Eds.; Wiley-VCH: Germany, 2008.

[24] Dadon, Z.; Wagner, N.; Ashkenasy, G. *Angew. Chem. Int. Ed.* **2008,** *47,* 6128-6136.

[25] Otto, S. *Acc. Chem. Res.* **2012,** *45,* 2200-2210.

[26] Cougnon, F. B. L.; Sanders, J. K. M. *Acc. Chem. Res.* **2012,** *45,* 2211-2221.

[27] Miller, B. L. *Dynamic combinatorial chemistry in drug discovery, bioorganic chemistry, and materials science;* Wiley: Hoboken, 2012.

[28] Reek, J.; Otto, S. *Dynamic combinatorial chemistry;* Wiley-VCH: Weinheim, 2010.

[29] Ladame, S. *Org. Biomol. Chem.* **2008,** *6,* 219-226.

[30] Lehn, J. M. *Chem. Soc. Rev.* **2007,** *36,* 151-160.

[31] Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **2002,** *41,* 898-952.

[32] Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006,** *106,* 3652-3711.

[33] Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *J. Am. Chem. Soc.* **2000,** *122,* 12063-12064.

[34] Oparin, A. I. *The Origin of Life;* Dover Publications: New York, 1953.

[35] Oparin, A. I. *The Origin of Life on the Earth;* Pergamon: Oxford, 1959.

[36] Maggini, L.; Bonifazi, D. *Chem. Soc. Rev.* **2012,** *41,* 211–241.

[37] Otto, S. *Acc. Chem. Res.* **2012,** *45,* 2200-2210.

[38] Black, S. P.; Wood, D. M.; Schwarz, F. B.; Ronson, T. K.; Holstein, J. J.; Stefankiewicz, A. R.; Schalley, C. A.; Sanders, J. K. M.; Nitschke, J. R. *Chem. Sci.* **2016,** *7,* 2614-2620.

[39] Wu, J.; Leung, K. C.-F.; Stoddart, J. F. *PNAS* **2007,** *104,* 17266-17271.

[40] Belowich, M. E.; Stoddart, J. F. *Chem. Soc. Rev.* **2012,** *41,* 2003-2024.

[41] Orgel, L. E. *Acc. Chem. Res.* **1995,** *28,* 109-118.

[42] Pross, A.; Pascal, R. *Open Biol.* **2013,** *3,* 120190.

[43] Robertson, A.; Sinclair, A. J.; Philp, D. *Chem. Soc. Rev.* **2000,** *29,* 141-152.

[44] Paul, N.; Joyce, G. *Curr. Opin. Chem. Biol.* **2004,** *8,* 634-639.

[45] Patzke, V.; von Kiedrowski, G. *Arkivoc* **2007,** *5,* 293-310.

[46] von Kiedrowski, G. *Angew. Chem. Int. Ed.* **1986,** *25,* 932-935.

[47] Bag, B. G.; von Kiedrowski, G. *Pure Appl. Chem.* **1996,** *68,* 2145.

[48] Issac, R.; Ham, Y. W.; Chmielewski, J. *Curr. Opin. Struct. Biol.* **2001,** *11,* 458.

[49] Lee, D. H.; Severin, K.; Ghadiri, M. R. *Curr. Opin. Chem. Biol.* **1997,** *1,* 491-496.

[50] Li, X.; Chmielewski, J. *Org. Biomol. Chem.* **2003,** *1,* 901.

[51] Paul, N.; Joyce, G. F. *Curr. Biol.* **2003,** *13,* R46.

[52] Wintner, E. A.; Rebek, J. *Acta Chem. Scand.* **1992,** *50,* 469-485.

[53] Rubinov, B.; Wagner, N.; Matmor, M.; Regev, O.; Ashkenasy, N.; Ashkenasy, G. *ACS Nano* **2012,** *6,* 7893-7901.

[54] Kassianidis, E.; Philp, D. *Chem. Commun.* **2006,** *0,* 4072-4074.

[55] Pearson, R. J.; Kassianidis, E.; Z., S. A. M.; Philp, D. *Chem.-Eur. J.* **2006,** *12,* 6829-6840.

[56] Kassianidis, E.; Pearson, R. J.; Philp, D. *Chem.-Eur. J.* **2006,** *12,* 8798-8812.

[57] Kassianidis, E.; Pearson, R. J.; Philp, D. *Org. Lett.* **2005,** *7,* 3833-3836.

[58] Pearson, R. J.; Kassianidis, E.; Z., S. A. M.; Philp, D. *Org. Biomol. Chem.* **2004,** *2,* 3434-3441.

[59] Pearson, R. J.; Kassianidis, E.; Philp, D. *Tetrahedron Lett.* **2004,** *45,* 4777-4780.

[60] Allen, V. C.; Philp, D. *Org. Lett.* **2001,** *3,* 777-780.

[61] Kassianidis, E.; Philp, D. *Angew. Chem. Int. Ed.* **2006,** *45,* 6344-6348.

[62] Robertson, C. C.; Mackenzie, H. W.; Kosikova, T.; Philp, D. *J. Am. Chem. Soc.* **2018,** *140,* 6832-6841.

[63] Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. *Nature* **1996,** *382,* 525-528.

[64] Severin, K.; Lee, D. H.; Martinez, J. A.; Ghadiri, M. R. *Chem.-Eur. J.* **1997,** *3,* 1017-1024.

[65] Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J. P.; Otto, S. *Science* **2010,** *327,* 1502-1506.

[66] Malakoutikhah, M.; Peyralans, J. J.-P.; Colomb-Delsuc, M.; Fanlo-Virgos, H.; Stuart, M. C. A.; Otto, S. *J. Am. Chem. Soc.* **2013,** *135,* 18406-18417.

[67] Pascal, R.; Pross, A. *J. Syst. Chem.* **2014,** *5,* 3.

[68] Pross, A. *J. Syst. Chem.* **2011,** *2,* 1.

[69] Belousov, B. P. *Collection of abstracts on radiation medicine* **1959,** *147,* 145.

[70] Zhabotinsky, A. M. *Biophysics* **1963,** *9,* 306-311.

[71] Debnath, S.; Roy, S.; Ulijn, R. V. *J. Am. Chem. Soc.* **2013,** *135,* 16789-16792.

[72] Boekhoven, J.; Brizard, A.; Kowlgi, K.; Koper, G.; Eelkema, Rienk andvan Esch, J. *Angew. Chem. Int. Ed. 49,* 4825-4828.

[73] Ilday, S.; Makey, G.; Akguc, G. B.; Yavuz, Ö.; Tokel, O.; Pavlov, I.; Gülseren, O.; Ilday, F. Ö. *Nature Comm.* **2017,** *8,* 14942.

[74] Bull, J. J.; Meyers, L. A.; Lachmann, M. *PLOS Comp. Biol.* **2005,** *1,* 450-460.

[75] Nowak, M. A. *Trends Ecol. Evol.* **1992,** *7,* 118-121.

[76] Sadownik, J. W.; Mattia, E.; Nowak, P.; Otto, S. *Nat. Chem.* **2016,** *8,* 264 - 269.